

CES Discussion Paper 10.03
Center for Economic Studies
K.U.Leuven

Product Differentiation on a Platform: the Informative and Persuasive Role of Advertising*

Dries De Smet[†] and Patrick Van Cayseele[‡]

January 30, 2010

Abstract

Both sides of a two-sided market are usually modeled as markets without product differentiation. Often however, it will be profit maximizing to differentiate one or two sides in two or more types. In a simple theoretical model, inspired by Yellow Pages, we show that this decision crucially depends on the appreciation of these differentiated types by the other side. We argue that this consists of two parts: first, a preference for informative advertisement by users and second, the effect of persuasive advertisements on users. The relation between both effects drives the monopolist decision to engage in product differentiation.

We test this conceptual framework in an empirical investigation of Yellow Pages. We find that Yellow Pages publishers offer large ads even though users don't value them at all. The economic rationale for this is that each advertisement type contributes directly (by the price paid for it) and indirectly (by increased usage) to revenues. Large ads are mainly set for this direct contribution, small ads for this indirect contribution. If a platform can choose the size, it will make the size difference between small and large ads as large as possible, in order to attract as much users as possible, but also to induce self selection among advertisers.

Keywords: two-sided markets, product differentiation, Yellow Pages, advertising

JEL Codes: D42, L12, L86

1 Introduction

Two-sided markets are mostly presented in a standard way: one or more platforms, two different demand sides and all members of each side join the platform receiving the same value proposition. But why not focus on product differentiation in these markets, i.e. why could members of each side not join the platform in a different way? Often, a platform can obtain a higher profit if it differentiates one or two sides in two or more types.

*We are grateful for helpful discussions with Peter Van der Hallen and seminar participants at the Media & Communication Conference of the AEA in Paris, 2007 and the Monetary & Information Economics Workshop in Leuven (2007 and 2009).

[†]dries.desmet@econ.kuleuven.be, Center of Economic Studies, K.U.Leuven (corresponding author)

[‡]patrick.vancayseele@econ.kuleuven.be, Center of Economic Studies, K.U.Leuven; Faculty of Economics and Econometrics, University of Amsterdam; LICOS Centre for Institutions and Economic Performance, K.U.Leuven

Take the example of Yellow Pages. A publisher is faced with two distinctly different demands: users searching for a good or service and firms announcing their goods or service to potential users. Both sides value each other's presence at the platform. A publisher can choose to offer an undifferentiated product, i.e. each side joins the platform in the same way. This is not what we observe in reality. Advertisers can select one out of many possibilities to join the platform: there are several advertisement types, differentiated by size, color and graphical elements.

We set up a simple model, presented as a Yellow Pages model for the ease of using that terminology, but can easily be applied to other markets. This model contains product differentiation on one side of the market: advertisers can choose a small or a large ad. These small and large ads are differentiated on two dimensions: their informative and persuasive content. We assume that large ads are more persuasive than small ads, i.e. if a user opens the directory, it is likelier that he will contact a firm that placed a large ad. On the informative side, we leave it open whether small ads or large ads are perceived as the more informative. One can argue that large ads contain more information, but they also contain more noise. Therefore the informative difference can go in both directions. The monopoly platform is a quantity setter. In his decision on both the quantity of small and large ads, he faces two constraints: quantities cannot be negative and he cannot place more quantities than there are firms. If none of these constraints is binding, he offers both types and both are generating profit. But if the non-negativity constraint is binding, he decides to offer only one type. This replicates the setup of many theoretical papers and some real life examples where no differentiation takes place. The platform also can decide to include every firm, by giving the smallest ads away for free.

Which result prevails is crucially determined by how these types are welcomed by each side of the market. When the users decide whether to join the platform or not, it is the informative role that matters. The persuasive role comes in only *after* one side has decided to join the platform. Once they are on the platform, it is not the information that counts, but the persuasion of the different types. The other side of the market, the advertisers, choose under which type they join the platform and are affected by both effects: the informative role determines how many users join; the persuasive role tells how large their chance is to be chosen once the users have joined. Since an individual advertiser cannot affect the aggregate information of the platform, its decision is mainly based on the persuasion role.

Whether the platform decides to offer differentiated types, or only one single type, is determined by the interplay of both roles. If large ads contribute strongly to usage (information) while users see them similar to smaller ads (persuasion), then the platform will offer only large ads to advertisers. If large ads do not contribute to the use of the directory but users do look at them, then small and large ads will figure next to each other. If the persuasion difference is large, then every firm is on the platform and small ads will not be charged.

In the empirical investigation of Yellow Pages in Europe, we isolate the informative and persuasive role. We find that readers use the directory because it contains listings and small ads; but the number of large ads does not have an effect on readership. It then seems puzzling why 18% of the firms buys a large ad. Our theoretical model correctly predicts that even though no one likes large advertisements, the directory includes them

because they are an important source of revenues. These results naturally extend to other media platforms and other two-sided markets, such as shopping malls, job agencies, auction platforms and other advertisement-related platforms.

The remainder of the paper is organized as follows. First, we relate our approach with the literature. In Section 3, we introduce our theoretical model. We present a simple model and compute the optimal quantities of a profit maximizing monopolist. Further we discuss some extensions, such as the optimal quantities of a welfare maximizing monopolist, price discrimination and endogenous size choice. The proofs of all propositions are relegated to the Appendix. Section 5 describes the industry study. We collected data on Yellow Pages publishers in five small European countries and test which advertisements are valued by Yellow Pages readers. Further, we reconcile empirical findings with theory. Finally, we conclude in Section 6.

2 Related Literature

The double effect of the informative and persuasive role of advertising is relevant in markets with product differentiation and network externalities. If there is no product differentiation, then there are no types to compare. The presence of two distinctive sides of the market allows information and persuasion to have a double perception by the other side of the market. We discuss two-sidedness and product differentiation below and relate them to the literature. We illustrate both characteristics with the example of Yellow Pages, because this makes it easier to link the theoretical model to the empirical analysis in section 5. However, first we start with a short discussion on the informative versus persuasive role of advertising.

The difference between both roles has been subject of discussion in economics and marketing for years (for an overview, see Bagwell (2007)). This discussion focused on the competitive effects of both advertisements types, arguing that persuasive ads are anti-competitive (Kaldor 1950) and informative ads are pro-competitive (Telser 1964). We refrain from this discussion and do also not discuss whether advertisements change preferences. We only use the terminology to describe the potential effects of advertisements on users. On the one hand, they watch advertisements for their informative content, but on the other hand, they are also influenced by the persuasive content. Since the information effect plays a role at the level of the directory, i.e. in deciding whether to use the directory or not, other terms for this effect are *aggregate* or *composition* effect. The persuasive effect comes in at the page level, i.e. after one has decided to open the directory, and can therefore be labeled *relative* or *comparison* effect.

Two-Sidedness In two-sided markets, platforms bring together two distinctly different customers who value each other's presence on the platform. Examples are credit cards (bringing together merchants and customers), operating systems (software developers and end users), shopping malls (sellers and buyers) and media outlets (readers and advertisers). Yellow Pages can be seen as a particular type of a media outlet.

Issuers of Yellow Pages target two clearly distinct groups: users and advertisers. For users, Yellow Pages is a device to find businesses. Though there are a number of alternatives, e.g. word-of mouth or company websites, Yellow Pages are still extensively used

for their convenience. For advertisers, Yellow Pages is a device to reach potential clients. Especially for smaller firms who lack the possibility of direct marketing, Yellow Pages are a valuable device. It is clear that both sides are interdependent: advertisers value eyeballs (i.e. the looks their ads get) while the users of the Yellow Pages value information on the suppliers that figure in each book entry.

Rysman (2004) examines the Yellow Pages market in the United States. He uses data on 419 directories for the year 1996. He estimates a simultaneous model on both sides of the market. The conclusion of his research is that there is a statistically and economically significant positive feedback loop effect. The willingness to pay of advertisers increases if there are more users and usage increases if there are more advertisers in a directory. This creates the feedback loop between advertisers and users: more ads means more users which in turn increases the number of ads. This result favors a monopolistic market situation because it generates a larger welfare surplus through the internalization of the network effect. But, more providers reduce market power, and therefore they are also a force that generates welfare surplus. His empirical study reveals that the latter effect dominates, therefore, from a welfare viewpoint, a more competitive market is preferable. For the empirical model, we rely on the structural framework of Rysman (2004), but we do not investigate the competition issue because pronounced consolidation already has induced one operator per country in Europe.

Rochet & Tirole (2003) include both network theory and multi-product pricing to build their seminal two-sided market model. They investigate the price allocation between the two sides of the market under a (monopolistic) platform and under two (competitive) platforms. They discuss the implications of profit vs. welfare maximization. Armstrong (2006) focuses on a similar analysis as Rochet & Tirole (2003). In his theoretical model, he mainly stresses the single versus multi-homing behavior of the market. If there are several platforms, both demand types can decide to join just one or different platforms. A characteristic example in his article is when one side of the market single homes (=joins just one platform) whereas the other side multi-homes. To some extent, the newspaper market is such a market: most readers buy only one newspaper, but many advertisers buy advertisement space in several papers. Armstrong (2006) shows that the *bottleneck side* (the single-homing side) is cross-subsidized by the multi-homing side. In the Yellow Pages market, users pay no fee and are therefore subsidized, though that is not necessarily the consequence of single/multihoming behavior.

We start from a simple two-sided markets model, but we remove some features and add some other. We remove the bottleneck discussion by assuming that there is but one supplier. This is true for Yellow in many European countries. Also, we assume that one side doesn't pay to use the platform, which equally fits with reality in Yellow Pages since users don't pay. Modeling a null price is equivalent with a relative high elasticity for the reader side. To the two-sided markets model, we add product differentiation.

Product differentiation Product differentiation is a well-known strategy to distinguish a product from a competitor's product, or from the own product. There are two types of product differentiation. Horizontal differentiation serves other tastes but keeps the same quality (e.g. a white versus a red car). In our model, we focus on vertical product differentiation. Products then differ in quality, e.g. large ads are more interesting for

advertisers because they attract more viewers. In this sense, product differentiation can also be labeled quality differentiation.

Within platforms, product differentiation can take place at two levels: at the demand sides of the market and at the platform level. Product differentiation on the demand sides is implicitly assumed in almost all two-sided markets articles. Especially in two-sided markets where unique matching takes place, such as dating markets, employment agencies and real estate, the number of customers on the other side of the market would be irrelevant if all those customers would be undifferentiated, i.e. just one customer would suffice to have the best match possible. Though, the number of customers on the other side is appreciated for its variety, resulting in a potentially better match¹. Hagiu (2009) focuses on this implicitly assumed but often neglected aspect of variety. If the taste for variety among buyers is greater, then profits will be more extracted from the seller side. The reasoning behind this finding is that a strong taste for variety reduces the competition among sellers².

The second form of differentiation in two-sided markets happens at the platform level. We distinguish two types: differentiation between platforms and differentiation on platforms. Armstrong (2006) and Rochet & Tirole (2003) differentiate platforms on a Hotelling line, with accompanying transport costs for both demands of the platform. This generates a differentiation between the platforms, but not on the platform. All users are offered the same product, though it is differently valued by consumers due to the transport cost.

Our model tackles differentiation on the platform. The advertisement side is differentiated on the platform, but the main difference between our model and previous literature is that the other side of the market can directly choose between both types of advertisers on the same platform, i.e. they do not have to multihome to connect with both types.

This differentiation is common to all Yellow Pages editions. Even casual inspection of a directory shows that different types of advertisements figure next to each other. Some are large ads in color that contain a lot of persuasive elements. Other entries merely have the contact details of the business involved. Advertisers are offered a menu of advertisements and can freely choose the advertisement which maximizes their utility. The interesting questions here are how advertisers self select their ad type and to what extent usage matters in this decision. It is also interesting to investigate how and to what extent types of advertisement affect usage.

Busse & Rysman (2005) investigate the effect of competition on second-degree price discrimination in the Yellow Pages market. With roughly the same data set as Rysman (2004), they find that competition does not only affect the price level, but also the price curvature. Their results suggest that the price of large ads is more affected by competition: in regions with only one player, the largest ad is 13.51 times more expensive than the

¹Take the example of Rochet & Tirole (2003). The number of merchants determines the welfare of buyers because they have a benefit each time they pay with plastic. Note that this number would be irrelevant if all shops are considered as the same, since one shop would suffice to do all purchases.

²Galeotti & Moraga-González (2009) also explicitly model the issue that both sides are differentiated and search for the best match. They focus on the potentially detrimental effect of an increase of own side members. If the variety increases, then it is also likelier that there will be a closer competitor, therefore competition is harsher. If there is an outside option, e.g. selling directly without intermediation, then this might sometimes be beneficial. This own-side effect externality is often labeled as a congestion effect. For a further discussion see section 4.7

smallest ad, compared to 12.54 and 11.63 in regions with two or three players, respectively. Similarly, an additional competitor triggers a decrease in the price for a full-page ad of 12.8%, for smaller ads this is substantially less (double quarter-column ad and quarter-column ad: 7.7% resp. 5.9%). Busse & Rysman (2001) provide three reasons for this observation. First, since new entrants have a capital constraint, they may seek out the most profitable sales which are the largest ads. Second, large advertisers have greater bargaining power (Rochet & Stole 2002); therefore they reap the largest benefits from competition. Finally, given the existence of a feedback loop effect, directories with more larger advertisement can get higher usage³.

There are a number of theoretical papers that allow for quality differentiation on at least one side of the market. Vicens (2006) presents a model inspired by shopping malls. There are two types of shops, differentiated by quality. For the other side of the market, buyers, not only the volume but also the identity of the shops matters. The relative size of the quality effect versus the quantity effect determines whether the duopolistic platforms will both engage in attracting low or high quality shops. If the network effects are weak vis-à-vis the reputation effect, only one platform will be active in the market. Otherwise, the market structure will be duopolistic with two platforms that might be asymmetric in their attraction of only low types, only high types or both. We do not allow for competition in our model since it would probably blur the interesting insights that result from the potential difference between the informative and persuasive role. Another major difference with Vicens (2006) is that we allow the differentiated side to self select their types.

In Damiano & Li (2007), a monopoly matchmaker differentiates both sides, men and women, in many types by setting a schedule of prices. The matchmaker uses price discrimination to allow men and women to select their optimal type, which generates an efficient matching process. They argue that differentiation is better than the uniform pricing that is used in online dating markets. Price is used as a signal and mitigates the misrepresentation in markets with uniform pricing. Our model is similar to their model to some extent since we also model the platform as a monopolist⁴ and we also allow the differentiated side to self select their type. Though, we do not model the potential inference of quality by a price mechanism. This inference would be perfectly possible in Yellow Pages: companies where large advertisements might signal that they are more efficient than the others. On the other side, the better companies might also want to attract the better customers which would replicate the matchmaking of Damiano & Li (2007). We refrain from this signalling possibility for two reasons. First, if one side cannot be charged, as is the case in Yellow Pages, then it is hard to differentiate on this side. Therefore, all users will connect with the largest advertisers and hence differentiation will also disappear at the advertiser side. Our theoretical model confirms this presumption. Second, we allow for cases where large is not always better, given the negative externality it might impose on the whole directory.

³This last explanation contradicts with both our theoretical and empirical findings. If larger ads are more popular than smaller ads, then the directory will only publish large ads. Empirically we find a strong positive effect on usage from small ads, but no effect from large ads.

⁴Interestingly, in their accompanying paper Damiano & Li (2008), they compare the results of monopoly with competition. They find that while monopolistic match makers can use prices to sort high types from low types, their duopolistic counterparts are involved too much in price competition and therefore they are less efficient.

This key feature of our model is not incorporated in the assortative matching model of Damiano & Li (2007).

With the theoretical model, we contribute to the already extensive literature on two-sided markets, by adding product differentiation. Besides the effect of this differentiation on the differentiated side of the market (by inducing self selection), we explicitly investigate the effect of this differentiation on the other side of the market. We argue that the other side has a direct preference for one type over another and an indirect preference for the contribution of each type to the composition of the platform. It is exactly the potential difference between both effects that determines the equilibrium outcome. To our knowledge, we are the first authors that exploit this feature to investigate the effect of product differentiation in two-sided markets.

Our contribution is not only novel in a theoretical perspective. Empirical studies on Yellow Pages are scarce; and none of them contain all the features described above. Rysman (2004) investigates two-sidedness but excludes product differentiation. Busse & Rysman (2005) do take into account product differentiation but exclude readers (and therefore two-sidedness). In our industry study, we focus on both, though we are mostly interested in what readers like.

3 Theoretical Model

Our theoretical model is presented as a Yellow Pages model, but is more general and can easily be applied to other markets as well. We merely do so because of the ease of explanation.

3.1 Set-Up

The Yellow Pages industry is characterized by three players: advertisers, readers and publishers. Advertisers are retailers or companies that use the directory to connect with users in order to sell their products or services. Readers are consumers that use the directory to connect with advertisers in order to buy their products or services. The publisher is the platform that connects readers and advertisers.

Our analysis is focused on the last player, the platform. We assume that the platform is a quantity setter⁵ and investigate the quantity decision of a monopolistic platform: how many small and large ads will be published in the directory?

Readers Yellow Pages readers get the directory for free, but that does not mean that every potential reader will use it. Readers use the directory if it generates more utility than the outside option to search for a good or service. This can include contacting a company one already knows, calling a friend, or driving to a city and searching randomly for retailers. If we treat this outside option as an opportunity cost, the net utility of using the directory is equal to:

$$U_R = \max(r_1 q_1 + r_2 q_2 - k, 0) \quad (1)$$

⁵While publishers have a tradition of announcing list prices each year, the practice learns that they adjust these prices with discounts to accommodate the number of advertisers, i.e. list prices can be seen as maximum prices but can differ substantially from the real prices.

where q_1 are the small ads, q_2 are the large ads and k is the opportunity cost. The parameters r_1 and r_2 represent the valuation of readers for the informative contribution of each ad type. Readers are homogenous in these valuations but are heterogeneous in their opportunity cost. If there is a mass one of potential readers and the opportunity cost k is uniformly distributed between 0 and 1, then the number of readers reads:

$$q_R = r_1 q_1 + r_2 q_2 \quad (2)$$

A directory without advertisements will have no readers.

Advertisers Firms, i.e. potential advertisers, value looks at their advertisement since they generate profit π from each look. The number of looks for each ad depends on the size⁶ of the ad s_i and the number of readers:

$$\begin{aligned} L_1 &= s_1 q_R \\ L_2 &= s_2 q_R \end{aligned} \quad (3)$$

We assume that large ads attract more readers than small ads ($s_2 > s_1$)⁷. Further we assume that each firm buys at most one ad. Contrary to readers, advertisers pay a price p to use the platform. The net benefit of an advertiser reads:

$$\pi L_i - p_i \quad (4)$$

with $i = 1, 2$ the type of ads. Advertisers are of mass one and are heterogeneous in their profit per look π which is uniformly distributed between 0 and 1. Define $\hat{\pi} \equiv \frac{p_2 - p_1}{L_2 - L_1}$ to be the profit level where firms are indifferent between both ad types, and $\hat{\pi}_1 \equiv \frac{p_1}{L_1}$ is the lowest profit for which the firm's net benefit is positive. Then

$$\begin{aligned} q_1 &= \hat{\pi} - \hat{\pi}_1 \\ q_2 &= 1 - \hat{\pi} \end{aligned} \quad (5)$$

Since we model the platform as a quantity setting monopolist, we transform the demand functions to inverse demand functions:

$$\begin{aligned} p_1 &= (1 - q_1 - q_2)L_1 \\ p_2 &= (1 - q_2)L_2 - q_1 L_1 \end{aligned} \quad (6)$$

Publisher The platform maximizes its profit:

$$\Pi = p_1 q_1 + p_2 q_2 \quad (7)$$

⁶Both parameters s_1 and s_2 stand for size. We relate size directly to persuasion, i.e. the larger the size, the larger the persuasion effect. Throughout the remainder of the analysis, we use the term size, which can equally be read as persuasion. Note that it can also be interpreted as salience, as in the framework of Haan & Moraga-Gonzalez (2009). Salience is the chance that a certain firm is remembered. Similar to persuasion, it increases profit.

⁷This assumption can be made without loss of generality. If it would be the case that small ads attract more attention than large ads, then small ads can be relabeled q_2 and large ads q_1 .

To simplify the analysis, we assume that the platform has no costs, i.e. there is no cost difference between small and large ads⁸. The platform is a quantity setting monopolist, so it optimizes its profit with respect to the quantities q_1 and q_2 . While deciding on these two variables, a platform has to take into account the direct and indirect effects of quantity level and structure. The effect of an increase in quantity on profit depends on the decrease in prices. This problem is complicated by product differentiation. Another indirect effect is the effect of quantity on usage. An increase in the number of ads increases the number of readers and therefore increases the willingness to pay. This effect, the feedback loop effect, reduces the negative price effect.

3.2 Results for a Private Monopolist

In this section, we present the results of the model for a private monopolist. In order to clearly present these optimal decisions, we define three new parameters: $r = \frac{r_1}{r_2}$, $R = r_1 + r_2$ and $s = \frac{s_1}{s_2}$. Thus r is the relative appreciation of q_1 versus q_2 (i.e. the informative effect), whereas s is the relative attention-attraction of q_1 versus q_2 (i.e. the persuasive effect). We replace r_1 , r_2 and s_1 to obtain expressions that contain only r , R , s and s_2 ⁹. The platform's profit function is maximized subject to the constraints $q_1, q_2 \geq 0$ and $q_1 + q_2 \leq 1$.

Proposition 3.1 *There are four possible outcomes for a monopoly platform:*

1. *Large ads only platform*
2. *Small ads only platform*
3. *Full coverage platform (i.e. everyone advertises)*
4. *Differentiated platform (i.e. small and large ads offered at a positive price)*

The solutions depend on the relation between r and s , but not on R and r_2 . These results will prevail, respectively, when:

1. $r < s$
2. $\frac{3}{4} < s < 1$ and $r \geq \frac{s}{4s-3}$
3. $s < \frac{1}{4}$ and $r \geq \frac{s-4s^2+2\sqrt{s^2(1-5s+4s^2)}}{1-4s}$
4. *otherwise*

From these solutions, we can easily deduct some possible scenarios.

Corollary 3.2 *r -scenarios*

- *if $r_1 < r_2$, then the platform always offers large advertisements*
- *if $r_2 < r_1$, then the platform always offers small advertisements*
- *if $r_2 = r_1$, then the platform always offers both advertisements*

⁸For a discussion of this assumption, see section 4.7

⁹We replace r_1 by $\frac{rR}{1+r}$, r_2 by $\frac{R}{1+r}$ and s_1 by ss_2 .

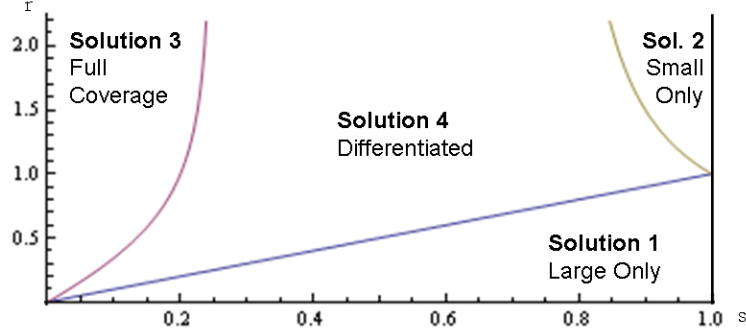


Figure 1: Different solutions in the r - s -space.

s-scenarios

- if $s \leq \frac{3}{4}$, then the platform always offers large advertisements
- if $s \geq \frac{1}{4}$, then advertisements are always paid (i.e. never for free)

Figure 1 shows these scenarios graphically in the r - s -space. As can be seen, the boundaries between solutions 3 and 4, $r = \frac{s-4s^2+2\sqrt{s^2(1-5s+4s^2)}}{1-4s}$, and between solutions 4 and 2, $r = \frac{s}{4s-3}$, go asymptotically to $\frac{1}{4}$ and $\frac{3}{4}$.

Performing a comparative statics analysis, allows us to formulate three interdependent statements on the impact of R and s_2 , r and s on the equilibrium quantities and profit.

Proposition 3.3 • The levels of the parameters (e.g. R , s_2) affect the profit positively, but do not affect the optimal quantities.

- The proportion of r_1 to r_2 ($=r$) does affect the optimal quantities, but only in the full coverage and differentiated optimum. The bigger r_1 relative to r_2 , the higher the amount of small ads and the lower the amount of large ads.
- The proportion of s_1 to s_2 ($=s$) does affect the optimal quantities, but only in the differentiated optimum. If r_1 is substantially larger than r_2 , an increase in s increases q_1 and decreases q_2 . If r_1 is smaller than r_2 , an increase in s decreases q_1 and increases q_2 .

An important remark on these comparative statics is that the effect of r and s does not play only within these solutions, but also affects which solution is optimal. A change in r or s can induce a solution shift. This can be easily seen in figure 1.

4 Extensions to the Model

4.1 Negative Utility from Advertisements

In section 3, we implicitly assumed that r_1 and r_2 are positive though it is likely that in some markets this assumption is violated. Especially in media markets, advertisements

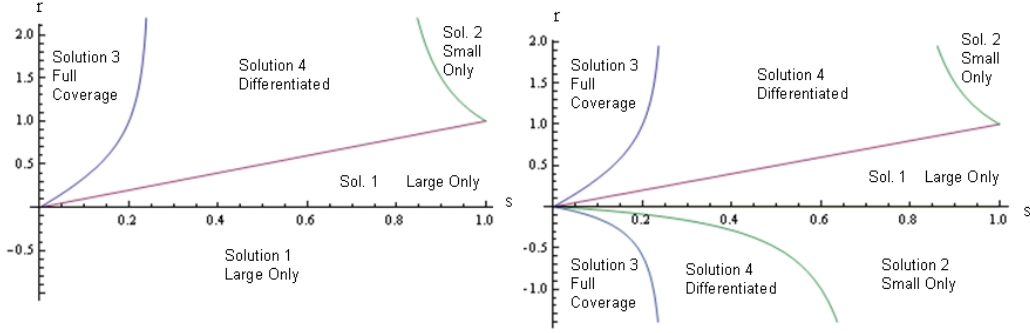


Figure 2: Different Solutions in the $r - s$ -space. Left panel: if $r < 0$, then $r_1 < 0$, right panel: if $r < 0$, then $r_2 < 0$

are often seen as a nuisance (Anderson & Coate (2005) and Peitz & Valletti (2008)). In Yellow Pages, this is less likely since advertisements contain relevant information for users, though it might be that some advertisement types do not contribute to usage, on the contrary, they decrease usage (see section 5.3 for a discussion).

In our model, it should hold that at least one of the advertisement types contributes to usage, i.e. either r_1 or r_2 is positive. Otherwise there would be no usage on the platform. But the model can perfectly cope with one negative parameter (either $r_1 < 0$ or $r_2 < 0$).

Proposition 4.1 • *If $r_1 < 0$ and $r_2 > 0$, then the platform only offers large advertisements.*

- *If $r_2 < 0$ and $r_1 > 0$, then the platform always offers small advertisements. The platform will also offer large advertisements if the size difference is large enough.*

This proposition is visualized in figure 2. In the left panel, below the x-axis, we plotted the case where $r_1 < 0$; on the right panel the case where $r_2 < 0$. As be seen from the graphs, the border $r < s$ extends to negative values of r as well in the case where $r_1 < 0$. Since small ads are less profitable and are disliked by users, they do not appear any more in the directory. In the case $r_2 < 0$, things are somewhat more complicated. The platform will keep out the large advertisements only if s is large; more precisely if $r > \frac{s}{4s-3}$ (and for all r -values if $r < 0$ and $\frac{3}{4} < s < 1$). Small ads will be offered for free if $|r| \geq \frac{s-4s^2 \pm 2\sqrt{s^2(1-5s+4s^2)}}{1-4s}$.

4.2 Endogenizing r (as a function of s_1 and s_2)

The persuasive effect s and the informative effect r were treated independently in section 3, though one can argue that they are related. In our Yellow Pages example, we argued that users might more easily choose a large advertisement when they compare both types. Though when deciding on using the directory or not, users might prefer smaller ads because these ads use less space for the information provided. Therefore, it makes the directory

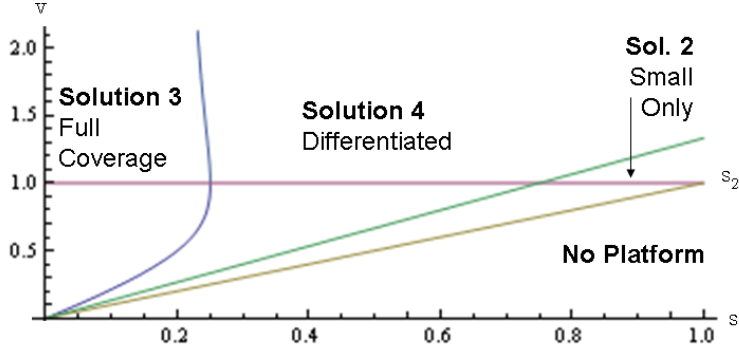


Figure 3: Different solutions in the v - s -space.

easier to handle and provides a better overview. We linked the persuasive effect directly to size. One can argue that the informative preference is also related to the size. After all, it is exactly the size of the advertisements that often blurs the informational value of an ad. Therefore, we can change usage equation 2 in order to include this nuisance effect of size.

$$q_R = v(q_1 + q_2) - n(s_1 q_1 + s_2 q_2) \quad (8)$$

The first term, with parameter v , captures the taste for variety, i.e. the pure network externality of adding another advertiser to the platform. The second term, with parameter n , captures the nuisance of the thickness of the directory. To simplify the analysis, we normalize n to 1. Note that we can rewrite equation 8 as $q_R = (v - s_1)q_1 + (v - s_2)q_2$, i.e. $r_1 = v - s_1$ and $r_2 = v - s_2$. The ratio r is equal to $\frac{v - s s_2}{v - s_2}$. Since $s_1 < s_2$, $r_1 > r_2$ always holds. Therefore, we can use the right panel of figure 2 to analyse the evolution of the equilibrium r -ratio.

Proposition 4.2 • If $v \leq s_1$, there will be no platform since there are no users.

- In all other cases, the platform will offer small ads. If $s_1 < v \leq \frac{4s_1}{3}$, then the platform offers only small ads. If $v > \frac{4s^2 s_2 - 2\sqrt{s^2(1-5s+4s^2)}s_2^2}{-1+5s}$, then the platform implements the full coverage solution. If $\frac{4s_1}{3} < v \leq \frac{4s^2 s_2 - 2\sqrt{s^2(1-5s+4s^2)}s_2^2}{-1+5s}$, then the platform is differentiated.

In figure 3, the outcomes are shown in function of v and s . The straight line is the size of the large advertisement s_2 . Three out of the four solution of the model in section 3 persist in this adapted model. In addition, if the taste for variety is not sufficiently large, then there will be no platform on the market. The claim that the absolute size of the advertisements does not matter in the choice of the quantities does not hold anymore.

Proposition 4.3 If size acts as a nuisance factor to the informational value of the platform, then an increase in s_2 will lead to an increase in small advertisements and a decrease in large advertisements.

This proposition holds for solutions 3 and 4 (full coverage and differentiated platform). In the case where only one ad type is offered, none of the parameters affect the solution. Note that an increase in s_2 can also induce a shift between the solutions and makes it likelier that there will be no platform in the market.

If size is not seen as a nuisance to the composition, but as a value added, the results change drastically. We can write the usage function as $q_R = v(q_1 + q_2) + n(s_1q_1 + s_2q_2)$. With n normalized to 1, the r of section 3 now reads as $r = \frac{v+s_1}{v+s_2}$ or $r = \frac{v+s_2}{v+s_2}$.

Proposition 4.4 *If size generates a positive effect on usage, then the platform always offers both types. If the size difference is large enough, then the small advertisement is offered for free. (If $v = 0$, then only large ads will be offered.)*

Since $r = \frac{v+s_2}{v+s_2}$ and $0 \leq s < 1$, for every combination between $v \geq 0$ and $s_2 \geq 0$, it holds that $s \leq r < 1$. If we have a look at figure 1 again, then we see that the only possible solutions are the full coverage and the differentiated solution.

4.3 Endogenous Size Decision

In our analysis, we treated size as exogenous, which gave us a good insight in the interplay between r and s . In reality, however, platforms can choose the size of the advertisements. Since we assumed that there are no costs in the model, platforms will set the level parameter s_2 of the advertisements as high as possible. It is more interesting to investigate the choice of s (for a given s_2). Except for the special cases $r = 0$ or $r = \infty$, the profit maximizing monopolist sets $s = 0$, i.e. it reduces the persuasive role of the small advertiser as much as possible. The profit maximizing monopolist implements solution 3, i.e. it offers readers a directory with all potential advertisers ($q_1 + q_2 = 1$). It offers two possibilities to readers: a large ad that attracts readers, or a small ad that attracts no readers. The latter is offered for free¹⁰.

Proposition 4.5 *If a profit maximizing monopolist can choose the size of the advertisements, it makes the large advertisement as large as possible and the small advertisement as small as possible. It offers readers a directory with all potential advertisers.*

This proposition has an important corollary in the debate whether a government should impose a universal service provision constraint. A universal service provision exists in the White Pages (people's directory)¹¹. It says that everyone has the right to be included in the directory for free¹². Such a constraint is not present for the Yellow Pages, though in

¹⁰Note that in solution 3, q_1 is always offered for free, even if $s > 0$.

¹¹There are three generally accepted telephone directories: White Pages, Yellow Pages and Grey Pages. The first type contains an alphabetical list of persons, with address and telephone number. The directory is divided in regions. Yellow Pages is synonymous with a business directory. It classifies firms by their business type or goods or services provided. Grey Pages are less known. These are so-called reverse telephone directories where one can browse the numbers and find the associated customer details. The latter was mostly used by emergency services, phone companies, law enforcement, and public libraries. All types are found online nowadays; it is likely that the less profitable printed White and Grey Pages will vanish first.

¹²See Directive 2002/22/EC of the European Parliament and of the Council of 7 March 2002 on universal service and users' rights relating to electronic communications networks and services (Universal Service Directive).

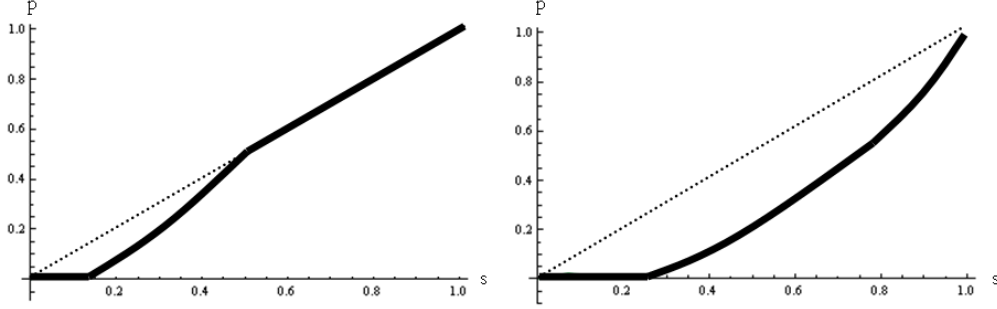


Figure 4: Relative prices ($p = \frac{p_1}{p_2}$). Left panel: $r=0.5$, right panel: $r=1000$

every country we examined, publishers include firms for free. This basic listing is as small as possible (one line) and contains only address and telephone number.

Corollary 4.6 *A universal service provision constraint is not necessary in Yellow Pages, since a profit maximizing publisher will always choose to open the directory for all potential advertisers.*

The main reason why it might be necessary to implement a universal service provision in White Pages, but not in Yellow Pages, is the distribution of profits per view. For a people's directory, this profit is nearly zero, while businesses can easier generate profit from being contacted. In other words, in White Pages, there is not enough advertisers that can cross-subsidize the less profitable types (see also the discussion on price discrimination in section 4.4).

If we add more advertisement categories in our model, as is the case in actual Yellow Pages, then the likely result would be that the largest advertisement is as large as possible, the smallest as small as possible. This is similar to the quality degradation models of Mussa & Rosen (1978), Maskin & Riley (1984) and Besanko, Donnenfeld & White (1988). In these models, a monopolist deteriorates the quality offered to the groups with the lowest willingness to pay for quality. If a regulator judges a mere listing as a product with too low quality, it can still implement regulatory corrections, such as minimum quality standards though this is not always welfare improving (Besanko, Donnenfeld & White 1988).

4.4 Price discrimination

If an advertisement of size one costs one euro, then we would expect that an advertisement of twice the size would cost less than two euro. If this does not hold, every advertiser can buy two advertisements and obtain the same effect. This phenomenon is well-described in microeconomics and explains why the bundle costs less than the sum of the individual parts. Busse & Rysman (2005) find the existence of this second degree price discrimination and link it to competition. If there are more platforms in the market, then the price discrimination is larger, i.e. large ads are relatively cheaper vis-à-vis small ads.

Using the model, we can have a look at the price ratio in function of the size ratio. We note that the absolute size (s_2) has no influence on prices. Prices are only determined by the size difference (s) and the information difference (r).

Proposition 4.7 *If both types are offered, then large advertisers always pay more per view than small advertisements. There is a cross-subsidization from large advertisers to small advertisers.*

The reason behind this cross-subsidization is that, in the case two types are offered, the relative contribution of small advertisements to usage is larger than their visibility ($r > s$).

Figure 4 shows the curvature, under $r < 1$ (left panel) and $r > 1$ (right panel). On the horizontal axes is the ratio of size ($s = \frac{s_1}{s_2}$). On the vertical axes, you find $p = \frac{p_1}{p_2}$, the ratio of prices. If there is second degree price discrimination, we would expect a price ratio above the 45 degree line. In both situations, we observe the reverse: if the size difference is maximal (s =small), then $p_1 = 0$ and therefore $p = 0$. The ratio gradually increases towards 1 when the size is the same.

One of the reasons why we don't find price discrimination is that we impose that advertisers buy at most one ad. Hence it is not possible to substitute a large ad by several small ads. Multiple advertisements are seldom observed in Yellow Pages, though this does not tell anything about the possibility. In sum, our model abstracts from the second degree price discrimination observed in Yellow Pages. On the contrary, the effects we investigate predict relative higher prices for the large ads.

4.5 Results for a Ramsey Planner

We contrast the optimal quantities of the private monopoly with a public monopoly platform that maximizes welfare under a break-even constraint. Therefore, it takes into account the welfare of readers and advertisers. The welfare of the readers is equal to the sum of the net utilities (see equation 1). The welfare of the advertisers is also the sum of the net utilities and consists of two parts: small advertisers and large advertisers (see equation 4). α and β are parameters that present the importance of advertiser welfare versus reader welfare. The objective function of the Ramsey planner reads:

$$W = \alpha W_A + \beta W_R \quad (9)$$

We impose the same restrictions as in the private monopoly case: $q_1, q_2 \geq 0$ and $q_1 + q_2 \leq 1$.

Proposition 4.8 *A welfare maximizing monopolist will always offer its advertisements for free and fully cover the market, though it offers only one type. The platform will offer only small ads if $r < \rho \equiv \frac{s_2\alpha + R\beta}{ss_2\alpha + R\beta}$; only large ads if $r \geq \rho$.*

Note that if the Ramsey planner cares only about readers ($\alpha = 0$), then $\rho = 1$. If he cares only about advertisers ($\beta = 0$), then $\rho = \frac{1}{s}$. If $r < 1$ or $r > \frac{1}{s}$, then there is no conflicting interest between readers and advertisers: the optimal solution is the same. But if $1 < r < \frac{1}{s}$, then the Ramsey planner has to disappoint either the readers or the advertisers. This can be seen in figure 5.

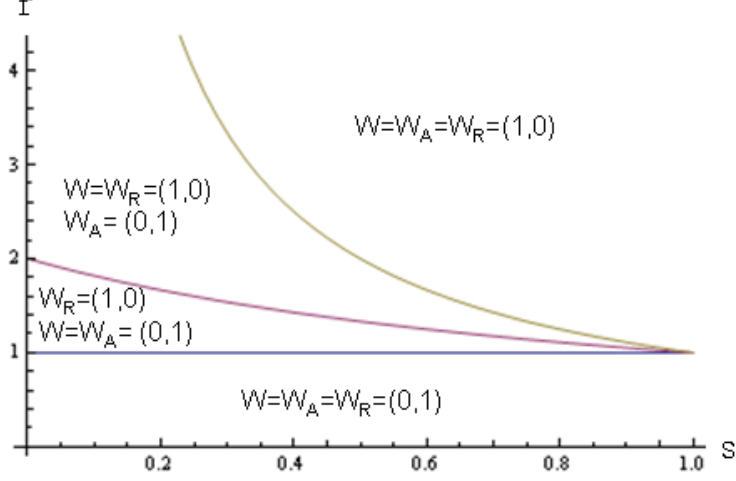


Figure 5: Optimal welfare solutions in the r - s -space. Between brackets, the optimal values for q_1 and q_2 are shown.

If a Ramsey planner can choose the size endogenously (see section 4.3, then he chooses $s = 0$ if $r > \rho \equiv \frac{s_2\alpha + R\beta}{ss_2\alpha + R\beta}$ and $s = 1$ otherwise. In other words, $s = 0$ is implemented if $q_2 = 1$, $s = 1$ if $q_1 = 1$.

4.6 Competition

We assumed the market structure is a monopoly, i.e. there is a single platform. We can examine what happens if two or more platforms enter the market within the context of the framework introduced.

Suppose that users singlehome, i.e. they use at most one platform to perform their searches. Competition will maximize the utility of the users. Therefore, platforms will offer only one ad type and give it away for free. If there exist an infinitesimally small entry cost, then the incumbent monopoly stays a monopoly but it makes no longer profits¹³.

Proposition 4.9 *If users singlehome and an infinitesimally small entry cost exist, then the incumbent monopolist becomes a contested monopolist. He will only offer the advertisement type that readers like most and will provide it for free.*

This proposition is partly driven by the fact that users are homogenous (except for their opportunity cost). Therefore, they all choose the platform with the largest gross utility. This result can change if heterogeneous users are introduced.

Note that competition causes the contested monopolist to implement solutions that are similar to the welfare solutions.

Corollary 4.10 *If the monopoly is contested, then it implements the optimal quantities that prevail under welfare optimizations. This does not hold if $1 < r < \frac{s_2\alpha + R\beta}{ss_2\alpha + R\beta}$. In this case, it offers the optimal welfare solution for readers only, but not for total welfare.*

¹³If there is no entry cost, results depend on how users decide using a certain platform or not when utility is exactly the same. If the market is split evenly among the platforms, there will be several platforms that all offer exactly the same directory.

This corollary can easily be explained by observing figure 5. Since we assume that readers singlehome, competition is focused on this side of the market. Since readers pay no price, the only option is to maximize welfare for them, leading the monopolist to implement always the reader welfare solution. The only conflict zone of reader welfare and total welfare is when $1 < r < \frac{s_2\alpha + R\beta}{ss_2\alpha + R\beta}$. Again, this result can be altered too, e.g. when singlehoming on the advertisement side is introduced.

4.7 Further Research

In this section, we discuss some future avenues that might improve the model we introduced above. We also suggest examples where our framework might apply.

Costs If a firm produces two goods with different quality, it is common to assume that production costs will be different. Normally, it will be costlier to produce the high quality product. Therefore, costs can also be one of the drivers of the platform decision to offer one or two types, and in which quantities. We excluded costs by setting them equal to zero for two reasons. First, we want to know the platform decision on quantities in a market with readers that hold double preferences and advertisers who self select their type. Introducing costs would add little to the model; but, it makes it more difficult to disentangle the reader and advertiser effects. Second, introducing costs complicates the mathematical computation and would make the model intractable. It is possible to solve a model with costs numerically, but this does not add novel insights from what we could see.

Congestion The visibility of ads remains unchanged if the number of ads increases. This means that there is no congestion effect or push away effect of other advertisers. It is unlikely that this holds in reality. We would expect that the chance of a look is reduced by an increase in advertisements¹⁴. This extension captures a negative own side effect.

The two-sided market literature towards congestion, or own-side effects, is mixed. Some articles ignore congestion, such as Rochet & Tirole (2003). They assume that both sides of the market always interact once, irrespective the number of agents on each side. This means that a potential congestion effect is absent. Other authors put congestion at the heart of their exposition. Belleflamme & Toulemonde (2009) investigate how negative intra-group externalities influence the potential of a for-profit platform to enter successfully in a market where the two sides are already trading on a non-profit platform. The for-profit platform enters by subsidizing one side and charging the other¹⁵. They show that entry is only possible if this externality is intermediate. Church & Gandal (1992) argue that software providers weigh the network effect (cross-group externality) with the competition effect of more agents on their side of the market, generating a lower chance of being seen

¹⁴This congestion effect is widely documented in the advertisement literature. As Comanor & Wilson (1974) put it: “To the extent that the advertising of others creates noise in the market, one must shout louder to be heard, so that the effectiveness of each advertising message declines as the aggregate volume of industry advertising increases.”

¹⁵While most articles do not allow for negative prices, subsidizing one side and charging the other side is a strategy very well suited for two-sided platforms. Caillaud & Jullien (2003) discuss these divide & conquer strategies and the efficiency of the resulting equilibria.

or chosen. The same question pops up in the empirical paper of Tucker & Zhang (2008): should platforms announce the number of sellers or not? On the one hand, a high number of sellers signals that there are also a high number of buyers (cross-group externality). On the other hand, it signals harsher competition for these buyers (congestion or own-side effect).

Theoretically such a congestion effect can easily be operationalized in our model by making the look functions L_1 and L_2 dependent on the quantities. The problem is similar to adding costs: by adding two or four additional parameters the model becomes intractable.

Other Applications Another interesting avenue for further research is to apply the framework to other markets. These markets should have the following in common with Yellow Pages. There should be two distinctly different demand sides that value each other's presence. One side can access the platform under different forms and it is important that they can self-select their type. The other side values one type over another in the direct comparison, but potentially values them differently on their contribution to the platform.

One example is a shopping street. Consider a shopping street as a platform that connects buyers with shopkeepers. Assume now that shops are differentiated by their display windows. Shops with large windows are more interesting to shopkeepers because they have a larger chance to be visited. It is a priori unclear whether buyers prefer small or large windows. Buyers can have a double preference for the size of the windows. On the one hand, a shop with a large window is more attractive than a shop with a small window (persuasion effect). On the other hand, if a buyer considers the entirety of shops, i.e. the shopping street, the large windows might be relatively annoying because they reduce the overall visibility and make the shopping street considerably longer (aggregation effect). Similar to the Yellow Pages example, the equilibrium number of shops with large and small windows can be described by parameters r and s , at least if there is one platform that controls all the shops in a certain area¹⁶.

Another example might be a dating event. Suppose that men and women are looking for a partner. Women are not charged. Men are charged but they can choose out of two types: they can appear at the event in suit or casual wear. If they compare them directly, women prefer men in suits over casually dressed men. But when they consider to go to the dating event or not, they might prefer an event with more casual men, because it means that they don't have to dress formally themselves. Again, the potential difference between direct comparison and contribution to the entirety of the platform might drive the decision of the platform to differentiate or not.

5 An Empirical Investigation into the Yellow Pages Industry

Since our results crucially depend on the relation between the informative and persuasive characteristics (r and s), it is interesting to document them for the industry that drives the

¹⁶While this is not so likely in an older city, it might be true in a shopping mall or newly developed area. For the implications of monopoly versus dispersed ownership of shops, see Nocke, Peitz & Stahl (2007).

theoretical model. Therefore, we measure r and reconcile the findings with the theoretical model.

First, we describe the main characteristics of the industry and the data. Then we discuss the identification of the two sides of the market and we present the results. We conclude with a comparison of our empirical and theoretical results.

5.1 Industry Characteristics

When a printer in Cheyenne, Wyoming, United States, ran out of white paper, yellow paper was used instead (Kane, Anzovin & Podell 2006). Invented by accident in 1883, Yellow Pages became a universal expression for telephone directories. Yellow Pages are still big business today. The total Yellow Pages revenues worldwide are estimated at 30.9 billion US dollar in 2008 (Kelsey 2009). Yellow Pages attract 5.2 % of the global advertising market and employ 82300 people, nearly half of them are sales representatives.

Print media are in turmoil everywhere in the world, and also the print Yellow Pages do not escape the crisis. Yellow Pages companies hold out relatively well in 2008, with revenues declining only 2.3% vis-à-vis 2007, but share prices nosedived at the end of the first decennium of the third millennium. If we compare the shares of the beginning of 2010 with the beginning of 2005, Seat (Italy) lost 99.7% of its price. Yell (UK) lost 91% and Pages Jaunes (France) 56%.

The reason for this decline, besides worldwide financial problems, is clear cut: the internet. In 2008, Yellow Pages companies got 85% of their revenues from their print division and 15% from online outlets. It is likely that the print edition will vanish, though the decline is slowed by two facts¹⁷. First, Yellow Pages are still seen as more reliable and extensive than online search engines. Second, firms are loyal to Yellow Pages, helped by an extensive sales force.

While the online market is characterized by many players, most European markets for print Yellow Pages are monopolies. In Europe, 70% of the countries has only one player in print¹⁸. Some countries have two publishers. These countries include the bigger countries (UK, Spain), Scandinavian countries (Sweden, Denmark, Finland) and alpine countries (Austria, Switzerland). A notable exception is Germany which has more than twenty active companies. We can expect further consolidation as was recently the case in the Netherlands (merger approved in 2008)¹⁹.

All Yellow Pages worldwide have the same appeal. Yellow Pages classify every business in different categories according to the goods or services it sells, such as plumbers, lawyers and swimming pool builders. Within each category, there are three main types of advertisements. The largest group is the free listings, which contains the basic information such as address and telephone number. The second largest group is the in column advertisements: small ads that fit into a directory column and add company details such as a

¹⁷The Economist, *Dial I for internet*, May 22nd 2008

¹⁸Data collected from the EADP website (EADP.org). EADP is the European Association of Directory and Database Publishers and coordinates most European Directory Publishers. We counted for each country the number of publishers in the category Telecommunication Directories that publish print directories. For most countries, the market structure is not monopolistic if one considers the market for internet Yellow Pages.

¹⁹NMa Decision 6246, case European Directories – Truvo Nederland

Table 1: Variables: summary statistics

Variable	Mean	Std.
Quantity Persuasive Ad	20937	10955
Quantity Informational Ad	96805	67842
Quantity Free Ad	657706	694509
Price Persuasive Ad	1622	1092
Price Informational Ad	310	199
Usage Penetration	0.40	0.12
Internet Penetration	0.29	0.17
Circulation	5807407	2423109
Population Density	213	137
Inflation (Index=100 in 2000)	101	8
Education	0.61	0.20
Income per Capita	26038	5511

url, opening hours or brand logo. The third type is the large ads. These advertisements span more than one column and contain besides additional information also persuasive content such as pictures and slogans.

Below, we describe the data, discuss the relevance of the advertisement types and present the results of our empirical study on Yellow Pages.

5.2 Data

To test the mutual effect of usage and advertisements in Yellow Pages, the first thing we need are data on usage and ads. Data on usage and advertisements are rare, even for those who advertise. Therefore, we contacted several companies to collect data.

We received book level data on advertising, i.e. quantities, list prices and real prices. The latter is important because it is common practice to give a discount off the price. This creates a gap between list prices and real prices. Since advertisers decide on real prices and not on list prices, we use real prices in the investigation.

Considering usage, we have book level data on distribution, but surveys on book usage are only available at the country level. To approximate the number of users, we multiply the percentage of people frequently using the book with the circulation in a certain area. Because the lack of detailed usage data, we carry out our analysis on country level and not on book level.

We also have data on country characteristics which are used as control variables. These data are obtained from international sources, such as IMF, OECD and Eurostat.

The sample includes 5 European countries. Those countries are relatively homogenous in population and market structure. It is dominated by countries with a single market player. In two countries, there is some competition, e.g. from local city directories, but we treat those countries as monopolies too. The earliest data point is 1995, the latest 2006. Since the data is compiled from different sources, we often lack data for the complete time span; i.e. the data set is an unbalanced panel. The statistics of the variables are summarized in table 1.

5.3 Identification

Reader Side Most users pick up the Yellow Pages to find a particular company or a particular good or service. If the directory is used to find a company which is already known to the user, then this is labeled as known search. If the directory is used to find new suppliers, then this is unknown search. The distinction between both types of look-ups can have a substantial impact on usage.

At one extreme, one might argue that the users only value “raw” information: the name and phone number of each supplier, be it classified by the category of the good or service that is supplied. In such a setting, it even might be the case that the users would pay for receiving a well classified directory that alphabetically lists all the suppliers in a certain category. This would be the case when each consumer is in a satisfactory or even optimal relationship with a certain supplier. The directory then is an ideal instrument for retrieving the coordinates of the particular supplier a consumer wants to patronize. But it merely serves the “administrative” purpose of an organized inventory of one’s business contacts. Large colored ads floating around then could disturb the user that is only interested in finding the coordinates of his trusted supplier. Hence these large ads might reduce the attractiveness of the directory.

At the other extreme, the large flashy ads in the directory serve to persuade consumers that have no relationship with a supplier yet, or look for change. These ads aim at the starting of a relationship with the particular supplier that uses the directory for this purpose. In this setting, larger and flashy ads may convey information that this type of user is keen for. It could separate the good suppliers from the bad in a signaling environment: the most efficient suppliers who have substantial turnover can pay for the larger ads, while the less efficient or inexperienced suppliers can not. The “nuisance” then comes from the small entries that merely provide for contact data. Since they signal no quality, they are not looked at and hence redundant. They could be dismissed entirely by the phone directory provider, wouldn’t it be for the obligation that he needs to list the phone numbers of all businesses in the area in the directory, for compliance with regulations²⁰.

In the case of the first extreme, people only use the directory as an index: only the company name and number suffices. In the other extreme, people use the directory to find out the reputation of a company: only persuasive ads count. But perhaps there is also a category in between: informative ads. Users do not search only for the telephone number, but also for more information about a company, i.e. fax number, e-mail or web address, opening hours or mobile number. Besides, it can also be interesting to have additional information about the activities of the company. If one is are looking for a replacement of one’s central heating, when looking at the category central heating, from a simple list of companies and telephone numbers, one can not find out whether a supplier is specialized in gas heaters or oil-fired central heating. An informative ad can provide more information.

We can model Yellow Pages as a usage generating market. By choosing the right amounts of quantities, a platform manages to create look-ups for the advertisers. One

²⁰Note that these regulations in Europe only include White Pages, not Yellow Pages. See the discussion in section 4.3.

potential production function is a Cobb-Douglas function.

$$U = BQ_0^{\beta_0}Q_1^{\beta_1}Q_2^{\beta_2}\prod_i control_i^{\beta_i} \quad (10)$$

If we take logs, then we obtain:

$$\ln(U) = \beta + \beta_0 \ln Q_0 + \beta_1 \ln Q_1 + \beta_2 \ln Q_2 + \sum \beta_i \ln control_i \quad (11)$$

We will estimate this equation in 5.4.

Usage is captured by the percentage of people answering yes on the survey question “Have you used the printed Yellow Pages last month?”. These are considered regular users of the directory. To count the number of persuasive ads, we add up the larger advertisements with logo or graphic. The informational ads are smaller ads that still fit into one column. These ads are in the within category alphabetical list of companies. Large persuasive ads are always accompanied by a regular entry in the alphabetical list, with a reference to the ad (e.g. see also advertisement on previous page). We don’t have data on the free listings, i.e. what companies get if they don’t advertise, but we can approach them by subtracting the total number of ads of the total number of companies in a country.

Advertiser Side Rysman (2004) treats ads as a homogeneous service and explains the amount or volume of advertising chosen by an individual business, but not the type. Equilibrium is reached when two offsetting network effects keep each other in balance. On the one hand, more advertising leads to increased usage of the directory, i.e. users like ads in Yellow Pages unlike what seems to be the case in other media. And since advertisers like eyeballs, they buy more advertising as usage increases. This is the positive feedback loop that links the two sides of the market. A countervailing force exists because of the negative network effect of congestion that takes place when too much advertising crowds a category: an overwhelming number of large ads for plumbers running over several pages of a directory is not likely the medium that still another plumber will choose to list his services in. The two effects taken together lead to directories in which not every business decides to buy ad space.

As noticed above however, the type of ad that is requested by advertisers can be quite different and hence this can result in a different impact regarding the network effects just mentioned. Large ads may attract more potential customers, but are perhaps annoying for users. At the same time they trigger the possibility that another business’ ad becomes unnoticed or that trade is diverted rather than created.

The small ads on the other hand probably convince fewer potential clients. But because they contain additional information in a condensed and surveyable way, they probably contribute more to the usage of the directory. By this, they might inflict a positive externality on the other advertisers of the directory. At the same time they might steal less business from the larger ads. In this respect, they might be complementary to larger ads, and the directory that has one large and two small ads might be a better product than the directory of the same size that is composed of two large ads.

It is not hard to understand that the providers of directories will take these differences into account. The result will be inter alia a different pricing strategy for each type of ad.

These pricing strategies will take into account that the contribution to usage of each type of ad is different and the effect of the ad on reaching potential customers. The congestion effect might be circumvented to some extent by including different types of ads that do not compete for the attention of the user. This leads to the claim that the different types of ads might be heterogeneous services.

Following Rysman (2004), we model the advertising site as a Cobb-Douglas function. Similar to the usage production function, it contains the major ad type and control variables. Besides, the price also depends on the number of users U .

$$p_j = A Q_0^{\alpha_0} Q_1^{\alpha_1} Q_2^{\alpha_2} U^{\alpha_3} \prod_i control_i^{\alpha_i} \quad (12)$$

If we take logs, then we obtain:

$$\ln p_j = \alpha + \alpha_0 \ln Q_0 + \alpha_1 \ln Q_1 + \alpha_2 \ln Q_2 + \alpha_3 \ln U + \sum \alpha_i \ln control_i \quad (13)$$

We will estimate this equation in 5.4.

In this price equation, the quantities are the same as in the usage equation. As explained above we prefer real prices above rate card prices. For usage, i.e. the number of consumers, we have no direct data. Therefore, we approach this number by multiplying the percentage of people regularly using the directory with the circulation. This may lead to an underestimation of the number of users, because directories are distributed to families which consist usually of more than one person. Robustness checks show though that multiplying this usage figure with average family size does not change the results below.

5.4 Results

Reader side Do users value a directory mostly for its listings, small ads or large ads? In this section we test this question and present the results.

Because we work with an unbalanced panel, consisting of 5 different countries, we use a least squares estimator with fixed effects. These estimation technique allows to control possible characteristics of particular countries - even without measuring them, as long as those characteristics do not change over time.

The number of observations is quite limited, though our results appear to be quite robust to other specifications of the model (see Appendix). We apply the three stage least squares methodology on these equations, which allows us to estimate a system of equations. Rysman (2004) proposed instrumental variables in his study on the Yellow Pages in the United States. In the usage equation, the advertisement level is instrumented by the number of people covered by a directory. In the advertisement equation, he instruments usage by the people that recently moved. For advertisements he uses the earnings level in a county, because this approximates the hourly wage and can be seen as a cost shifter. We do not apply these instruments because we have only a limited number of observations and we do not have the necessary data to instrument.

As can be seen from table 3, the coefficients of the quantities of the small advertisements and the free listings are significantly differing from zero. The coefficient of small ads is quite large: increasing the number of small ads with 1 percent increases the number of users with 2 percent. Large ads seem to have no effect on usage.

Table 2: Results of 3SLS regression

	Usage	Price Small Ads	Price Large Ads
Quantity Index Ads	0.391 (3.86)***	-0.003 -0.07	0.227 (5.23)***
Quantity Small Ads	2.166 (2.72)***	-0.764 (2.47)**	2.138 (7.06)***
Quantity Large Ads	0.03 -0.04	-0.141 -0.62	-1.288 (5.56)***
Internet	0.087 -0.21	0.368 (2.82)***	-0.021 -0.16
Income per Capita	-0.805 -0.98	0.364 -1.29	-0.34 -1.19
Population Density	8.457 (1.85)*		
Usage		0.196 (1.91)*	-0.089 -0.87
CPI		-0.004 -0.74	0.026 (4.53)***
Constant	-51.764 (2.29)**	9.184 (2.93)***	-5.779 (1.83)*
Observations	24	24	24
R-squared			

Absolute value of z statistics in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

These results indicate that readers value most the small ads in the book, but listings (=only name and telephone number) are also appreciated. Large advertisements are considered as annoying and do not affect usage. While large advertisements can be interesting for advertisers because they attract a lot of eyeballs, users are not interested in the additional (persuasive) elements.

Advertiser side In the same table 3, the results for the price equation are given. Since price, quantity and usage are given in logs, we can easily interpret the coefficients. As expected, both ad categories have a negative and significant own elasticity.

Higher usage increases the willingness to pay in the small ads category, though it is not significantly different from zero in the large ad category. The same phenomenon pops up if we consider the cross term (quantity of other advertising category). In the large ads price equation, the quantity of small ads has a positive and significant effect on the price. This means that more small ads increase the willingness to pay for large ads. The amount of large ads has no effect on the price of small ads. The quantity of free ads affects the price of large ads positively.

The fact that internet penetration has a significantly positive effect on the price of small ads, might be related to the fact that the internet is not a substitute, but a complement for additional information, e.g. it would be of no use to add a mail and website address if no users had internet access. As explained above, internet might rather be a substitute for large ads, because it serves the same needs: in depth search for a product or service.

Reconciling the empirical and the theoretical results With our empirical results in mind, it is interesting to look back to the theoretical model. The most interesting

application of these results comes from the usage equation. If we add a third advertisement category to the theoretical model, say q_0 (which is associated with $s_0 < s_1$, then we get results for estimations of r_0 , r_1 and r_2 . Point estimates for these parameters are 0.39, 2.16 and 0.03 respectively. If we ignore r_0 for a moment, the estimates of r_1 and r_2 can be merged to a r estimate. Since $r = \frac{r_1}{r_2}$, $r = 72$. One can wonder, if ratio is really that large, why is q_2 still in the directory? The answer can be seen in figure 1. With a high r , the result depends on the value of s . If s is smaller than 0.25, then q_1 is offered for free; if s is larger than 0.752, then only q_1 is offered. Between 0.25 and 0.75 both quantities are offered and a positive price occurs in the market.

Even if no one likes the large advertisements, Yellow Pages will still offer them as long as the difference in views is large enough. If our models extends to three categories, then if the listing is small enough, then it will be offered for free and the whole market is covered. If the large advertisement is substantially larger than the small advertisement, then all three advertisements will be offered. In most directories, the free listing is only one line, which fits with our prediction that, if platforms can choose the size difference, they will make it as large as possible.

We don't have size information for all observations, though rough estimations show that the size of the average small advertisement is between 0.04 and 0.11 of size of the average large advertisement. This is relatively small but still a lot larger than the listing. The average listing is between 0.002 and 0.006 of the size of a large advertisement. With the assumptions of our model, this size difference would lead to a solution where all advertisers are included and the smallest advertisement is given away for free. Our model does not incorporate the possibility of a third advertisement type, which probably explains why small advertisements are paid in reality.

It is harder to reconcile the estimation of the advertisement equation with the theoretical model. The reason is that we have not modeled the potential congestion effect in the directory. We have assumed that the number of rival advertisements have no influence on the visibility of the advertisement, or $\frac{ds_1}{dq_1} = \frac{ds_2}{dq_1} = \frac{ds_1}{dq_2} = \frac{ds_2}{dq_2} = 0$. What we do model is the self-selection effect. Therefore, an increase in the amount of small advertisement should not only decrease the price the small ads, but also of the large ads. We do find the negative effect of an increased quantity on the own price, but we do not find an effect on the cross price. A shift in large ads has no effect on small ads, while small ads have a *positive* effect on large ads. The theoretical model falls short in explaining this sign.

6 Conclusion

Yellow Pages readers do like small advertisements, but place no value at large advertisements. An empirical investigation on Yellow Pages in five European countries shows this quite clearly. It then might seem puzzling why 18% of the firms in our sample buy a large ad.

One of the reasons is given by the theoretical model we introduced: the discrepancy between what readers value when they compare ad types and when they look at the entirety of the platform. We relate this comparison to the persuasive role of advertisements and the valuation of the entire directory to the informative role. If large ads are sufficiently valued by users, then the platform will choose to offer only large ads to advertisers. If

the reverse holds, then small and large ads will be offered except when the persuasion difference between small and large ads is small. If the difference is large, then every potential advertiser is included in the directory and the small ads are offered for free.

If platforms can choose the size difference, then they maximize it and implement the solution with full coverage in the market. This resembles reality where every firm is included for free with a mentioning as small as possible; firms that want to attract more readers have to pay for more space. If the platform has no profit motives but wants to maximize welfare, then the platform implements a corner solution: every firm gets a small ad, or every firm gets a large ad; depending on the ratio of the parameters. When both are included, we find a cross-subsidization from large ads to small ads, which is quite different from what a price discrimination story would tell.

Future research might include prices and costs on both sides and own side externalities (congestion effect). It might also address competition between platforms. Further it would be interesting to apply this model and test the results in other markets than Yellow Pages, such as advertisement-related industries in general, shopping malls, dating events, job agencies and auction platforms.

Appendix

Proof of Proposition 3.1 If we solve equation 7 with respect to the constraints $q_1, q_2 \geq 0$ and $q_1 + q_2 \leq 1$, we find four optimal solutions:

	q_1^*	q_2^*
solution 1	0	$\frac{2}{3}$
solution 2	$\frac{2}{3}$	0
solution 3	$\frac{2-r-\sqrt{1-r+r^2}}{3(1-r)}$	$\frac{1-2r+\sqrt{1-r+r^2}}{3(1-r)}$
solution 4	$\frac{r(s-1)s+s^2-s^3+\sqrt{\diamond}}{6(1-s)s(r^2+s-2rs)}$	$\frac{3r^3(s-1)-8r^2(s-1)s+2s^2-2s^3-\sqrt{\diamond}+r(5s^3-3s^2-2s+\sqrt{\diamond})}{6(s-r)(1-s)s(r^2+s-2rs)}$

where $\diamond = (r-s)^2(s-1)s(-3r^2+6rs+s^2-4s)$. As can be seen in figure 6, solution 4, the differentiated optimum, will always be profit maximizing. Though, this solution is not always feasible. Since we have solved this model with Karush-Kuhn-Tucker conditions, we should check whether and when these conditions hold.

The only non-zero Lagrange operator related to the constraint $q_1 + q_2 \leq 1$, prevails in solution 3. Therefore, if the Lagrange operator, which is equal to $\frac{R(r^3(1-4s)-3(1+\sqrt{1-r+r^2})s+r^2(-4+\sqrt{1-r+r^2}+(7-4\sqrt{1-r+r^2})s)+r(1+\sqrt{1-r+r^2}+(2+5\sqrt{1-r+r^2})s))s^2}{9(-1+r)^2(1+r)}$, is positive, then solution 3 should be implemented and $q_1 + q_2$ will be equal to 1. The corresponding price p_1 is equal to 0, i.e. small advertisers don't pay to be included.

The condition $q_1 \geq 0$ is violated if $r \leq s$, in this case solution 1 will be implemented where $q_1 = 0$. The other non-negativity condition $q_2 \geq 0$ is violated if $r > \frac{s}{4s-3}$, in this case solution 2 will be implemented where $q_2 = 0$.

Proof of Proposition 3.3 To trace back the effects of R , s_2 , r and s , we perform a comparative statics analysis on the equilibrium quantities of the different solutions and on the profits under the different solutions.

First, we find that the derivatives under all solutions of quantity with respect to R and s_2 are zero: $\frac{dq_1}{dR} = \frac{dq_2}{dR} = \frac{dq_1}{ds_2} = \frac{dq_2}{ds_2} = 0$. If we derive the equilibrium profits Π^i (with

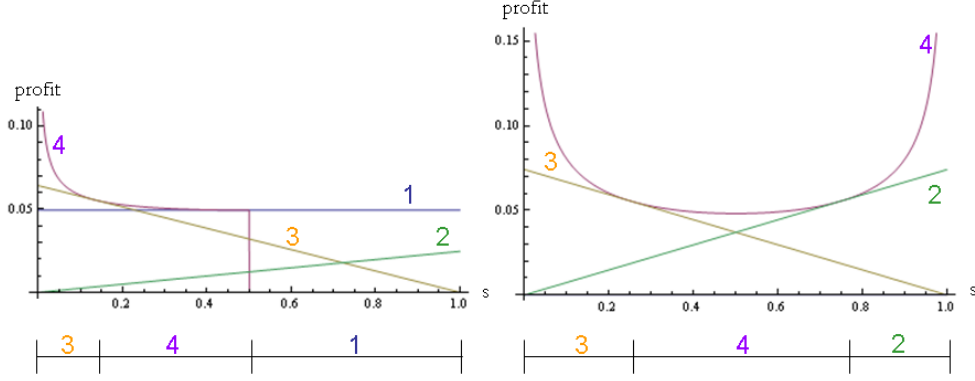


Figure 6: Platform profit. Left panel: $r=0.5$, right panel: $r=1000$

index i for the solution) towards R and s_2 , we obtain:

$$\begin{aligned}
 \frac{d\Pi^1}{dR} = \frac{d\Pi^1}{ds_2} &= \frac{4\alpha}{27(1+r)} \\
 \frac{d\Pi^2}{dR} = \frac{d\Pi^2}{ds_2} &= \frac{4r\alpha s}{27(1+r)} \\
 \frac{d\Pi^3}{dR} = \frac{d\Pi^3}{ds_2} &= \frac{(1-2r+\sqrt{1-r+r^2})(-1+2r^2-\sqrt{1-r+r^2}+2r(-1+\sqrt{1-r+r^2}))\alpha(-1+s)}{27(-1+r)^2(1+r)} \\
 \frac{d\Pi^4}{dR} = \frac{d\Pi^4}{ds_2} &= -\frac{\alpha(-9r^3(-1+s)s+3r^2(-9s^2+9s^3+\sqrt{(r-s)^2(-1+s)s(-3r^2+6rs+(-4+s)s)})}{108(1+r)(r-s)(-1+s)s(r^2+s-2rs)} \\
 &\quad -\frac{rs(-8s-11s^2+19s^3+6\sqrt{(r-s)^2(-1+s)s(-3r^2+6rs+(-4+s)s)})}{108(1+r)(r-s)(-1+s)s(r^2+s-2rs)} \\
 &\quad +\frac{s(-8s^2+7s^3+s^4+4\sqrt{(r-s)^2(-1+s)s(-3r^2+6rs+(-4+s)s)}-s\sqrt{(r-s)^2(-1+s)s(-3r^2+6rs+(-4+s)s)})}{108(1+r)(r-s)(-1+s)s(r^2+s-2rs)}
 \end{aligned}$$

with $\alpha = s_2$ in the case of $\frac{d\Pi}{dR}$ and $\alpha = R$ in the case of $\frac{d\Pi}{ds_2}$. If $r \geq 0$, $R > 0$, $s_2 > 0$ and $0 \leq s < 1$, then these derivatives are unambiguously positive. An increase in R or s_2 always increases profit.

Second, if we derive the optimal quantities to r , we obtain:

$$\begin{aligned}
 \frac{dq_1^1}{dr} &= 0 \\
 \frac{dq_1^2}{dr} &= 0 \\
 \frac{dq_1^3}{dr} &= \frac{-1-r+2\sqrt{1-r+r^2}}{6(-1+r)^2\sqrt{1-r+r^2}} \\
 \frac{dq_1^4}{dr} &= \frac{(r-s)\left(r(-1+s)s+s^2-s^3+\sqrt{(r-s)^2(-1+s)s(-3r^2+6rs+(-4+s)s)}\right)}{3(-1+s)s(r^2+s-2rs)^2} \\
 &\quad -\frac{-6r^3+18r^2s+4s^2+2s^3-2rs(2+7s)+\sqrt{(r-s)^2(-1+s)s(-3r^2+6rs+(-4+s)s)}}{6(r^2+s-2rs)\sqrt{(r-s)^2(-1+s)s(3r^2-6rs-(-4+s)s)}}
 \end{aligned}$$

Results for $\frac{dq_2}{dr}$ are similar, and, as can be expected in the full coverage case, $\frac{dq_2^3}{dr}$ is exactly same as $\frac{dq_1^3}{dr}$, but with the opposite sign. With the same restrictions as above, we can show that $\frac{dq_1}{dr} > 0$ and $\frac{dq_2}{dr} < 0$ in the differentiated and full coverage case.

Third, we can derive the optimal quantities towards s . It is easy to show that $\frac{dq_1^1}{ds} = \frac{dq_1^2}{ds} = \frac{dq_1^3}{ds} = 0$. The derivative $\frac{dq_1^4}{ds}$ is somewhat more complicated; and crosses

the horizontal axes once in the r -space, i.e. it can be both negative and positive, depending on the value of r .

We can summarize all the comparative statics as follows:

	$\frac{dq_1}{dr}$	$\frac{dq_1}{dR}$	$\frac{dq_1}{ds}$	$\frac{dq_1}{ds_2}$
solution 1	0	0	0	0
solution 2	0	0	0	0
solution 3	>	0	0	0
solution 4	>	0	< 0 if $r < \rho_1$ > 0 if $r > \rho_1$	0

	$\frac{dq_2}{dr}$	$\frac{dq_2}{dR}$	$\frac{dq_2}{ds}$	$\frac{dq_2}{ds_2}$
solution 1	0	0	0	0
solution 2	0	0	0	0
solution 3	<	0	0	0
solution 4	<	0	> 0 if $r < 1$ < 0 if $r > 1$	0

	$\frac{d\Pi}{dr}$	$\frac{d\Pi}{dR}$	$\frac{d\Pi}{ds}$	$\frac{d\Pi}{ds_2}$
solution 1	<	>	0	>
solution 2	>	>	>	>
solution 3	< 0 if $r < 1$ > 0 if $r > 1$	>	<	>
solution 4	< 0 if $r < \rho_2$ > 0 if $r > \rho_2$	>	< 0 if $r < \rho_3$ > 0 if $r > \rho_3$	>

where ρ_1 , ρ_2 and ρ_3 are parameter values above 1 where there is a switch in the comparative statics, i.e. ρ_1 is the value of r for which $\frac{dq_1^2}{ds} = 0$ and mutatis mutandis the same holds for ρ_2 and ρ_3 .

Proof of Proposition 4.1 The rationale behind this proposition is similar as proposition 3.1. It can be seen as an extension or rectification of this proposition. As in proposition 3.1, solution 4 is the unconstrained optimum of the problem. The same constraints apply: $q_1 + q_2 \leq 1$ and $q_1, q_2 \geq 0$.

In the case where $r_2 > 0$ and $r_1 < 0$, it suffices to investigate $q_1 \leq 0$ since the platform will never set $q_2 \leq 0$ or $q_1 + q_2 \leq 1$, even if that would be possible. The constraint $q_1 \leq 0$ is violated if $r \leq s$, also for values $r < 0$. Therefore, the analysis is not altered.

Things are somewhat more complicated in the case where $r_1 > 0$ and $r_2 < 0$. The constraint $q_1 \geq 0$ will never be violated since this would induce a no-usage platform. To check whether $q_1 + q_2 \leq 1$ is not violated in solution 4, we have to investigate in which range the Lagrange operator is positive. We find two r -regions where the assumptions is violated: $r \geq \frac{s-4s^2-2\sqrt{s^2(1-5s+4s^2)}}{1-4s}$ if $r > 0$ and $r \leq \frac{s-4s^2+2\sqrt{s^2(1-5s+4s^2)}}{1-4s}$ if $r < 0$. We also have to check whether $q_2 = \frac{3r^3(s-1)-8r^2(s-1)s+2s^2-2s^3-\sqrt{\diamond}+r(5s^3-3s^2-2s+\sqrt{\diamond})}{6(s-r)(1-s)s(r^2+s-2rs)} \geq 0$ (with $\diamond = (r-s)^2(s-1)s(-3r^2+6rs+s^2-4s)$). This is violated in two r -regions: $\frac{s}{4s-3}$ in both cases ($r < 0$ and $r > 0$) and for all r -values if $r < 0$ and $\frac{3}{4} < s < 1$.

Proof of Proposition 4.2 We solve the same model as in section 3, though we alter the usage equation. Since the new usage equation can be written in a similar way as the old usage equation, we obtain qualitatively the same results. Though, solution 1, where $(q_1, q_2) = (0, \frac{2}{3})$ is no optimal solution anymore. This solution can be replaced by solution 5, $(q_1, q_2) = (0, 0)$, which prevents that q_R would drop below zero.

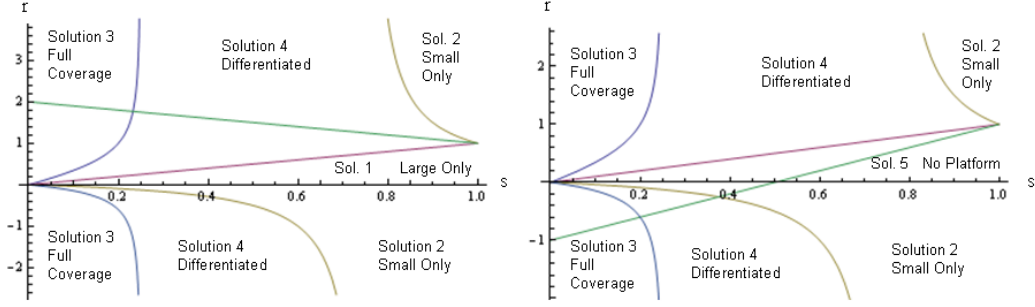


Figure 7: Different Solutions in the $r - s$ -space. Left panel: if $r < 0$, then $r_1 < 0, r_2 > 0$, right panel: if $r < 0$, then $r_1 > 0, r_2 < 0$

Since we can rewrite the old parameters r_1 and r_2 as $v - s_1$ and $v - s_2$, we can interpret the solutions in the $r - s$ -space. r is equal to $\frac{v - ss_2}{v - s_2}$, i.e. r is determined by s (in the graph) and v and s_2 (which are fixed in the graph). There are two cases that should be investigated. First, if $v > s_2$, then both quantities are valued positively in the usage equation. Second, if $v < s_2$, then at least one quantity is valued negatively. If $s_1 < v < s_2$, then small advertisements generate a positive effect on usage, large advertisements generate a negative effect. If $v < s_1$, then both advertisements generate a negative effect and it is optimal for the platform to offer no advertisements at all.

In figure 7, the optimal solution choice is given for $v > s_2$ (left panel) and $v < s_2$ (right panel). The $r(s)$ corresponding with the s -values is given by the straight green line. In both cases, the relative value of v versus s_2 determines the slope. If $v > s_2$, the larger v , the flatter this line. If $v < s_2$, the larger v , the steeper this line. As can be seen in the left panel, the r determination crosses solution 3 and 4 (for v set twice as large as s_2), but if v was smaller, then it would have crossed solution 2 as well. If we know the value s , then we can immediately derive from the graph which solution will prevail. The same holds for the case $v < s_2$. Note that if $r(s) > 0$, this indicates that both r_1 and r_2 are smaller than zero and that no advertisements are offered in optimum. If $r(s) < 0$, then the line always crosses the solution 2, 3 and 4. On the right panel, $r(s)$ is drawn for v half as large as s_2 .

Proof of Proposition 4.3 In solutions 3 and 4, the optimal quantities read:

$$\begin{array}{ll} \text{solution 3} & \frac{q_1^*}{(-2+s)s_2+v-\sqrt{(1-s+s^2)s_2^2-(1+s)s_2v+v^2}} \\ \text{solution 4} & \frac{s(s_2-v)v+\sqrt{sv^2(3v^2+s(4s_2^2-8s_2v+v^2))}}{6s(ss_2(s_2-2v)+v^2)} \end{array} \quad \begin{array}{ll} q_2^* & \frac{(-1+2s)s_2-v+\sqrt{(1-s+s^2)s_2^2-(1+s)s_2v+v^2}}{3(-1+s)s_2} \\ & \frac{ss_2(2s_2-5v)+3v^3-s_2\sqrt{sv^2(3v^2+s(4s_2^2-8s_2v+v^2))}}{6v(ss_2(s_2-2v)+v^2)} \end{array}$$

If we derive these quantities to s_2 , then $\frac{dq_1}{ds_2} > 0$ and $\frac{dq_2}{ds_2} < 0$, at least in the regions where these solutions prevail.

Proof of Proposition 4.4 We can draw the same graph as for proposition 4.2. The $r(s)$ corresponding with the s -values is given by the straight green line. This line always

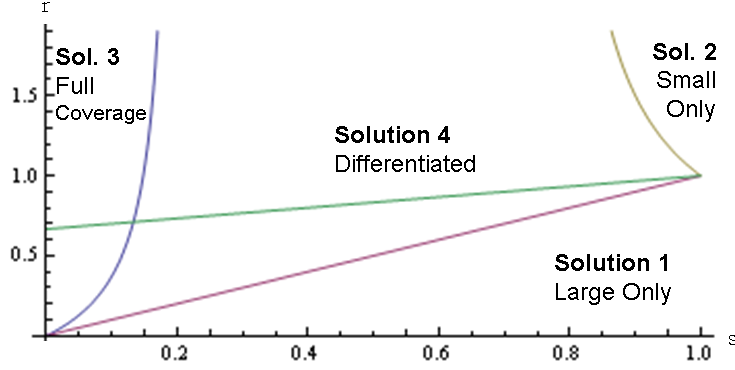


Figure 8: Different solutions in the r - s -space.

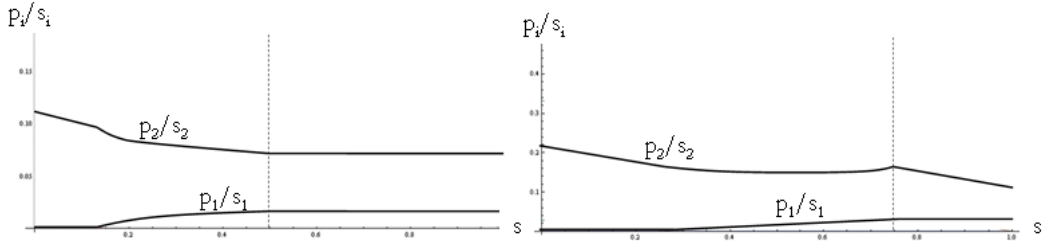


Figure 9: Price ratios. Left panel: $r=0.5$, right panel: $r=1000$

ends in point $(1,1)$. The starting point $(0, \frac{v}{v+s_2})$ depends on the values of v and s_2 and should lie between zero and one. If $v = 0$, then $r = s$ and we obtain solution 1, i.e. only large quantities are offered (see the proof of proposition 3.1. In graph 8, the $r(s)$ -line is shown for $v = 2$ and $s_2 = 1$.

Proof of Proposition 4.5 To find the solution of the private monopolist who decides on quantities and size (s), we maximize equation 7 with respect these variables. We find two solutions for this problem, which are equivalent to solutions 2 and 3 of the problem in section 3.2, supplemented with a size choice of $s = 1$ in solution 2 and $s = 0$ in solution 3. Further we find that solution 3 dominates for all values of r . The only exception is $r \rightarrow \infty$; then the platform is indifferent between both solutions. In sum, one can say that $s = 0$ is optimal for all values of r . Moreover, it is the only solution if $r_2 > 0$.

Proof of Proposition 4.7 To check whether the price per view is higher for large advertisements, one simply has to check whether $\frac{p_2}{s_2} > \frac{p_1}{s_1}$, for every solution in the relevant

parameter space. The ratios $\frac{p_1}{s_1}$ and $\frac{p_2}{s_2}$ are plotted in figure 9. On the left panel, $r = 0.5$, on the right panel $r = 1000$. The s -values on the left hand side of the dotted line are the relevant ones, because on the right hand side of this line only one type is offered in the market.

Proof of Proposition 4.8 We solve equation 9 with the same Karush-Kuhn-Tucker conditions as under the private monopolist, i.e. $q_1, q_2 \geq 0$ and $q_1 + q_2 \leq 1$. We obtain three groups of solutions: an unconstrained optimum, two solutions where either q_1 or q_2 is zero and two solutions where $q_1 + q_2 = 1$. We find further that only the Lagrange operator related to this last solution is always binding in the region $0 < s < 1$. Therefore, $(1,0)$ and $(0,1)$ are the only relevant solutions to our problem.

We obtain the critical $r = \rho \equiv \frac{s_2\alpha + R\beta}{ss_2\alpha + R\beta}$ by comparing the two welfare levels, under the first and second solution. Equivalently we can say that a Ramsey planner implements solution 1 if $s < \frac{s_2\alpha + R\beta - rR\beta}{rs_2\alpha}$ and solution 2 otherwise. The critical values for the monopolist maximizing the welfare of one side of the market only, can be found easily by setting α or β equal to zero.

Proof of Proposition 4.9 Suppose that a platform would offer a quantity combination that does not maximize reader welfare. Then another platform can offer a combination that slightly improves reader welfare and the first platform will have no users and hence no advertisers. Since there are no costs in our model, it is always feasible to offer all advertisements for free. If we introduce costs to our model, then the platform can never offer the solution where there is only one type and that type is free of charge.

Table 3: Results of 3SLS, OLS and SURE regression

	3SLS			OLS			SURE		
	Usage	Price Small Ads	Price Large Ads	Usage	Price Small Ads	Price Large Ads	Usage	Price Small Ads	Price Large Ads
Quantity Index Ads	0.391 (3.86)***	-0.003 -0.07	0.227 (5.23)***	0.391 (3.25)***	-0.023 -0.45	0.137 (3.02)***	0.394 (3.89)***	-0.02 -0.47	0.152 (4.13)***
Quantity Small Ads	2.166 (2.72)***	-0.764 (2.47)**	2.138 (7.06)***	2.166 (2.29)**	-0.851 (2.26)**	1.744 (5.29)***	2.23 (2.80)***	-0.832 (2.71)***	1.829 (6.82)***
Quantity Large Ads	0.03 -0.04	-0.141 -0.62	-1.288 (5.56)***	0.03 -0.04	-0.155 -0.57	-1.352 (5.71)***	0.006 -0.01	-0.155 -0.7	-1.354 (7.01)***
Internet	0.087 -0.21	0.368 (2.82)***	-0.021 -0.16	0.087 -0.18	0.373 (2.35)**	0 0	0.113 -0.28	0.371 (2.87)***	-0.009 -0.08
Income per Capita	-0.805 -0.98	0.364 -1.29	-0.34 -1.19	-0.805 -0.82	0.346 -1.07	-0.423 -1.5	-0.939 -1.14	0.341 -1.3	-0.446 (1.94)*
Population Density	8.457 (1.85)*			8.457 -1.56			9.611 (2.11)**		
Usage		0.196 (1.91)*	-0.089 -0.87		0.253 (3.07)***	0.17 (2.35)**		0.247 (3.66)***	0.141 (2.40)**
CPI		-0.004 -0.74	0.026 (4.53)***		-0.004 -0.64	0.027 (4.92)***		-0.004 -0.75	0.028 (6.25)***
Constant	-51.764 (2.29)**	9.184 (2.93)***	-5.779 (1.83)*	-51.764 (1.93)*	9.905 (2.63)**	-2.52 -0.76	-56.982 (2.53)**	9.776 (3.18)***	-3.104 -1.15
Observations	24	24	24	24	24	24	24	24	24
R-squared				0.76	0.96	0.94			

Absolute value of z statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

References

- Anderson, S. P. & Coate, S. (2005), ‘Market provision of broadcasting: A welfare analysis’, *Review of Economic Studies* **72**(4), 947–972.
- Armstrong, M. (2006), ‘Competition in two-sided markets’, *RAND Journal of Economics* **37**(3), 668–91.
- Bagwell, K. (2007), *The Economic Analysis of Advertising*, Vol. 3 of *Handbook of Industrial Organization*, Elsevier, chapter 28, pp. 1701–1844.
- Belleflamme, P. & Toulemonde, E. (2009), ‘Negative intra-group externalities in two-sided markets’, *International Economic Review* **50**(1), 245–272.
- Besanko, D., Donnenfeld, S. & White, L. J. (1988), ‘The multiproduct firm, quality choice, and regulation’, *Journal of Industrial Economics* **36**(4), 411–29.
- Busse, M. R. & Rysman, M. (2001), Competition and price discrimination in yellow pages advertising, Yale School of Management Working Papers ysm207, Yale School of Management.
- Busse, M. & Rysman, M. (2005), ‘Competition and price discrimination in yellow pages advertising’, *RAND Journal of Economics* **36**(2), 378–390.
- Caillaud, B. & Jullien, B. (2003), ‘Chicken and egg: Competition among intermediation service providers’, *RAND Journal of Economics* **34**(2), 309–28.
- Church, J. & Gandal, N. (1992), ‘Network effects, software provision, and standardization’, *Journal of Industrial Economics* **40**(1), 85–103.
- Comanor, W. S. & Wilson, T. A. (1974), *Advertising and Market Power*, Harvard University Press.
- Damiano, E. & Li, H. (2007), ‘Price discrimination and efficient matching’, *Economic Theory* **30**(2), 243–263.
- Damiano, E. & Li, H. (2008), ‘Competing matchmaking’, *Journal of the European Economic Association* **6**(4), 789–818.
- Galeotti, A. & Moraga-González, J. L. (2009), ‘Platform intermediation in a market for differentiated products’, *European Economic Review* **53**(4), 417–428.
- Haan, M. A. & Moraga-Gonzalez, J. L. (2009), Advertising for attention in a consumer search model, IESE Research Papers D/794, IESE Business School.
- Hagiu, A. (2009), ‘Two-sided platforms: Product variety and pricing structures’, *Journal of Economics & Management Strategy* **18**(4), 1011–1043.
- Kaldor, N. (1950), ‘The economic aspects of advertising’, *Review of Economic Studies* **18**(1), 1–27.

- Kane, J., Anzovin, S. & Podell, J. (2006), *Famous first facts: a record of first happenings, discoveries, and inventions in American history*, 6 edn, H.W. Wilson.
- Kelsey (2009), Global yellow pages 2009-2010, Technical report, Kelsey Group.
- Maskin, E. & Riley, J. (1984), ‘Monopoly with incomplete information’, *RAND Journal of Economics* **15**(2), 171–196.
- Mussa, M. & Rosen, S. (1978), ‘Monopoly and product quality’, *Journal of Economic Theory* **18**, 301–17.
- Nocke, V., Peitz, M. & Stahl, K. (2007), ‘Platform ownership’, *Journal of the European Economic Association* **5**(6), 1130–1160.
- Peitz, M. & Valletti, T. M. (2008), ‘Content and advertising in the media: Pay-tv versus free-to-air’, *International Journal of Industrial Organization* **26**(4), 949–965.
- Rochet, J.-C. & Stole, L. A. (2002), ‘Nonlinear pricing with random participation’, *Review of Economic Studies* **69**(1), 277–311.
- Rochet, J.-C. & Tirole, J. (2003), ‘Platform competition in two-sided markets’, *Journal of the European Economic Association* **1**(4), 990–1029.
- Rysman, M. (2004), ‘Competition between networks: A study of the market for yellow pages’, *Review of Economic Studies* **71**(2), 483–512.
- Telser, L. G. (1964), ‘Advertising and competition’, *Journal of Political Economy* **72**, 537.
- Tucker, C. & Zhang, J. (2008), Decomposing the congestion effect and the cross-platform effect in two-sided networks: A field experiment, Technical Report 8-12, NET Institute.
- Viezens, M. F. (2006), Two-sided platforms with endogenous quality differentiation, Economics Working Papers we061204, Universidad Carlos III, Departamento de Economía.

